237966

Preprint
UCRL-JC-139281

Artificial neural network solutions of slab-geometry neutron diffusion problems

P. S. Brantley

This article was submitted to
American Nuclear Society ANS/ENS 2000 International Winter Meeting & Embedded Topical Meetings, Washington, DC, November 12-16, 2000

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

June 12, 2000

ARTIFICIAL NEURAL NETWORK SOLUTIONS OF SLAB-GEOMETRY NEUTRON DIFFUSION PROBLEMS

Patrick S. Brantley
Lawrence Livermore National Laboratory
P.O. Box 808, L-023
Livermore, CA 94551

Artificial neural network (ANN) methods have been researched extensively within the nuclear community for applications in systems control, diagnostics, and signal processing. We consider here the use of multilayer perceptron ANNs[1] as an alternative to finite-difference and finite-element methods for obtaining solutions to neutron diffusion problems. This work is based on a method proposed by van Milligen et. al. to obtain solutions of the differential equations arising in plasma physics applications.[2] This ANN method has the potential advantage of yielding an accurate, differentiable approximation to the solution of diffusion problems at all points in the spatial domain.

We consider slab-geometry monoenergetic neutron diffusion problems with spatially-constant total and absorption cross sections, σ_t and σ_a, a constant interior source Q, and Marshak boundary conditions with left and right boundary incident currents, J_0^+ and J_R^-, respectively. Denoting the diffusion operator as $\mathcal{L}$ and the left boundary condition and right boundary condition operators as $\mathcal{L}_0$ and $\mathcal{L}_R$, respectively, we write the diffusion problem as

$$\mathcal{L}\phi\left(x\right) \equiv -\frac{1}{3\sigma_t}\frac{d^2}{dx^2}\phi\left(x\right) + \sigma_a\phi\left(x\right) = Q \quad , \quad 0 < x < R \quad , \tag{1}$$

$$\mathcal{L}_0\phi\left(0\right) \equiv \frac{1}{4}\phi\left(0\right) - \frac{1}{6\sigma_t}\frac{d}{dx}\phi\left(0\right) = J_0^+ \quad , \tag{2}$$

$$\mathcal{L}_R\phi\left(R\right) \equiv \frac{1}{4}\phi\left(R\right) + \frac{1}{6\sigma_t}\frac{d}{dx}\phi\left(R\right) = J_R^- \quad . \tag{3}$$

We approximate the scalar flux $\phi\left(x\right)$ using a multilayer perceptron ANN with one hidden layer. This type of network is capable of approximating any continuous function to an arbitrary accuracy if a sufficient number of hidden layer neurons are present.[1] The input to the ANN is the spatial variable x, and the activation of the j^{th} hidden layer neuron, $1 \leq j \leq J$, is given by

$$y_j = f\left(\alpha_j x + u_j\right) \quad . \tag{4}$$

Here α_j is the connection weight between the input neuron and the j^{th} hidden layer neuron, u_j is the so-called "bias term" of the j^{th} hidden layer neuron, and $f\left(y\right)$ is the binary sigmoid activation function given by $f(y) = [1+\exp(-y)]^{-1}$. The output of the neural network is given by

$$N = \sum_{j=1}^{J} \beta_j y_j + v \quad , \tag{5}$$

where β_j is the connection weight between the j^{th} hidden layer neuron and the output neuron, and v is the bias term of the output neuron. The connection weights and bias terms are typically initialized to small random values and are systematically adjusted during the training of the ANN.

The ANN is trained to approximate the solution of the diffusion problem by minimizing an objective function given by

$$E = \frac{\gamma}{2}\sum_{i=1}^{I}\left(\mathcal{L}N(x_i) - Q\right)^2 + \frac{\gamma_0}{2}\left(\mathcal{L}_0 N(0) - J_0^+\right)^2 + \frac{\gamma_R}{2}\left(\mathcal{L}_R N(R) - J_R^-\right)^2 \quad , \tag{6}$$

where I is the number of spatial training points in the interior of the slab and the γ terms are weights accounting for the relative magnitudes of the residual terms. The derivatives arising from the action of the diffusion and boundary condition operators on the neural network can be calculated analytically. The objective function given by Eq. (6) is a measure of how accurately the output of the ANN approximates the solution of the diffusion problem at the interior and boundary training points. This objective function is minimized by systematically adjusting the weights of the ANN using a minimization algorithm (e.g. conjugate gradient or quasi-Newton). The gradient of the objective function with respect to the ANN parameters (utilized by the minimization algorithm) can be calculated analytically.

We note that the training of the ANN requires minimization of a nonlinear objective function and can suffer from trapping in local minima. The accuracy with which the ANN approximates the neutron scalar flux depends on the initial choice of ANN weights (typically initialized to small random values). Certain initializations of the ANN weights lead to convergence to a local minima of the objective function and a corresponding inaccurate approximation to the scalar flux. One approach to solving this problem would be to incorporate a global optimization algorithm (e.g. simulated annealing) into the neural network training procedure. However, further research into the resolution of this problem is required.

We consider here two numerical problems solved in a domain $0 \le x \le 1$ with total cross section $\sigma_t = 10$ and absorption cross section $\sigma_a = 0.5$. Problem I has an interior source $Q = 0.5$ and vacuum boundaries. Problem II has no interior source, an incident current on the left edge of the slab of magnitude $J_0^+ = 0.5$, and a vacuum boundary on the right edge of the slab. In both cases, the ANN used five hidden layer neurons and was trained using eleven spatial training points (one point on each of the two boundaries and nine equally-spaced interior points).

In Table 1 we compare typical globally-converged results obtained using the ANN method to those obtained from a reference fine-mesh finite-difference solution of the diffusion problem. The error values tabulated are root mean squared relative errors in the neutron scalar flux at both the spatial training points (training error) and at 101 equally-spaced points across the slab (interpolation error). For both problems, the average training errors are significantly less than 1%. The average interpolation errors are larger than the average training errors, as expected, but exhibit reasonable accuracy for points in the spatial domain not used in the training of the ANN.

Table 1: RMS Relative Errors in Neutron Scalar Flux for Problems I and II

Problem	Training Error	Interpolation Error
I	2.065×10^{-3}	4.179×10^{-3}
II	1.033×10^{-3}	1.879×10^{-2}

This ANN method is not computationally efficient in one spatial dimension. However, the compact representation of the neutron scalar flux afforded by the ANN should be advantageous for multidimensional problems. The relative efficiency and accuracy of the ANN method compared to finite difference and finite element methods for multidimensional problems remains to be seen.

In conclusion, we have described the use of multilayer perceptron ANNs for the solution of monoenergetic slab-geometry neutron diffusion problems with Marshak boundary conditions. The ANN method produces a reasonably accurate approximation to the neutron scalar flux both at the spatial training points and at interpolation points within the spatial domain of the problem.

We are also investigating the use of ANNs for the solution of slab-geometry neutron transport problems, in which case the ANN has inputs for both the spatial and angular variables. The conceptual advantages to this approach as opposed to traditional discrete ordinates techniques are: 1) the spatial and angular variables are treated consistently; 2) complete freedom exists for the selection of training points in the spatial and angular domains. We plan to report on this work in the future.

Acknowledgments

This work was performed under the auspices of the U.S. Department of Energy by the University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

References

[1] L. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*, Prentice Hall, Upper Saddle River, New Jersey (1994).

[2] B.Ph. van Milligen, V. Tribaldos, and J.A. Jimenez, *Phys. Rev. Letters* **75**, 3594 (1995).